REVIEW OF PACIFIC COD MODELING

Ana M. Parma
Centro Nacional Patagónico
9120 Puerto Madryn, Chubut, Argentina

EXCECUTIVE SUMMARY

The increased recognition of the importance of ecological and management questions related to the spatial distribution of fish stocks have created a need to collect data at finer spatial resolutions and to develop modeling approaches for their analysis. Dr. Grant Thompson, the lead author of the Pacific cod assessment, has explored alternative assessment models to serve these needs. The Alaska Fisheries Science Center (AFSC) requested a review of the modeling framework developed by Dr. Thompson, which is being proposed for future use in the Bering Sea/Aleutian Islands and Gulf of Alaska Pacific Cod assessments.

The review took place at the AFSC in Seattle between July 30 and August 3 of 2001. The consultant met with the author and became familiar with the various subject areas involved in the proposed modeling framework, including assumptions about the population dynamics and spatial processes, the Kalman filter, the AD Model Builder implementation of the estimation model, and the development of risk-adverse fishing targets for management. Main findings and conclusions from the review were:

1. The new model offers an innovative and useful approach for the analysis of fish stock dynamics over time and space.
2. A number of simplifying assumptions related to growth, recruitment and selectivity have been made, which allow for parameterization of the model in terms of quantities of management interest, namely, fishing mortality at maximum sustainable yield (F_{MSY}) and its corresponding biomass level (B_{MSY}). These assumptions may or may not hold in different situations, so their adequacy should be evaluated for each specific application.
3. New results from decision theory have been derived to determine optimal risk-averse fishing targets in the presence of process variability, state estimation uncertainty, and parameter estimation uncertainty. The results provide quantitative rules for decreasing fishing targets as a function of these uncertainties. These rules can be very valuable to guide management decisions, especially in cases where direct evaluations of management strategies via simulations are not available.
4. The model uses survey and commercial CPUE to estimate changes in fish relative distribution over space and time. The use of commercial data may be problematic if fishing effort targets preferentially areas of high abundance, and those move over time. Preliminary simulations conducted during the review indicate that in such cases the estimated spatial distributions tend to smear the areas of high abundance over time and space. This problem will need to be further examined using simulation-estimation trials tailored to specific real case situations (*e.g*. Pacific cod).

5. The advantages of using spatially-structured stock assessments are clear when the population dynamics depend on spatial location, and catch and effort data are available at the model spatial resolution to estimate location-dependent model parameters. In the current model formulation, the dynamic processes (i.e. recruitment, mortality, growth) are location-independent and only the data on catch-per-unit-of-effort are spatially explicit. Consequently, the benefits of *simultaneously* modeling the stock dynamics and the spatial distribution are not obvious.
6. The new model could be readily extended to allow fishing mortalities to vary over space. In this way, it could make best use of spatially-explicit data on catch and effort when these data become available. Until that time, the two aspects –the spatial dynamics and the age-structured dynamics– might be more efficiently addressed separately. The pros and cons of such an approach could be explored with simulations.
7. The program used to estimate model parameters, coded using AD Model Builder, will need to be optimized to facilitate conduction of estimation-simulation trials using realistic dimensions. Specific suggestions are made to that end.
8. Beyond the context of the specific model reviewed, the lack of spatially explicit data is seen as a key limitation that will need to be resolved before spatially explicit approaches can be utilized to address management problems at small spatial resolutions.

BACKGROUND

The most recent assessment of Pacific cod was conducted in 1999 using the well-known Stock Synthesis assessment model. The specific implementation of Stock Synthesis for Pacific cod is based on estimating model parameters by fitting the model to data on total catch, survey biomass estimates, and size compositions for the survey and commercial catches. Age-composition data are not yet available but aging methods have now been developed that will make it possible to assemble a series of catch-at-age data in the near future. The assessment approach was subject to a review in 1997, after which some changes were introduced. However, the basic Synthesis framework is still in use.

The 1999 assessment was based on considering the Pacific cod stock in the Eastern Bering Sea and Aleutian Islands as a single management unit. The increased recognition of the importance of ecological and management questions related to the spatial distribution of fish stocks, not just their overall abundance, have created a need to collect data at finer spatial resolutions and to develop modeling approaches for their analysis. The case of Pacific cod is particularly important in this regard, as cod has been found to be an important prey item in the diet of Steller sea lions, an endangered species. This motivated Dr. Grant Thompson, the lead author of the Pacific cod assessment, to explore alternative approaches for conducting spatially-explicit stock assessments. He has now developed a space- and age-structured model with the intention of applying it to Pacific cod and probably other stocks in the near future.

SCOPE OF REVIEW

The Terms of Reference for this review (Appendix 1) reflect the interest of the Alaska Fishery Center to encourage the development of spatially explicit modeling approaches to address management needs related to critical habitat, predator-prey interactions and other ecosystem considerations. Specifically, the review focused on the new modeling approach proposed by Dr. Thompson. The model is fully developed and its performance is currently being evaluated using simulated data. The evaluation is still at an early stage and the model has not yet been applied to any particular stock. Thus, the focus of the review was on the actual model structure in terms of basic assumptions, without regard to any specific application. Data for Pacific cod were not available at the time of the review to address stock-specific concerns with regard to model assumptions, as discussed below.

DESCRIPTION OF REVIEW ACTIVITIES

The review took place at the AFSC in Seattle. Dr. Thompson and I met regularly during a five-day period from July 30 to August 3. The week prior to the review, the author provided me with background documentation and copies of computing files containing the mathematic derivations and computing code for simulations and parameter estimation (Appendix 2). The review started with a detailed presentation of the material by the author, followed by discussions of some preliminary evaluations of model performance done using simulations. After that initial presentation, I spent the rest of the week reviewing the documents and discussing with the author different aspects of the modeling approach, as reflected in this report. In many cases these discussions led to exploratory simulations and further work that Dr. Thompson tackled very effectively. The work he conducted over the course of the review facilitated reaching some conclusions and identifying areas that will need further examination in the future.

This report does not add anything substantive to the issues addressed during the meetings in Seattle, which I see as the main product of the review. The ideas and recommendations presented here are a result of fruitful exchanges and collaborative work between Dr. Thompson and myself. Of course, I take full responsibility for the views expressed below, which are not necessarily his. I am thankful for the opportunity to review this work, and appreciate the support given by Dr. Thompson and other scientist at the Center, which facilitated my work.

SUMMARY OF FINDINGS ABOUT THE NEW MODEL AND ITS IMPLEMENTATION

Overview

The population model reviewed is structured by age and space. The population dynamics model and the spatial model were first developed as stand-alone models, which were later combined into a single age-structured, spatially-explicit population model. The approach is innovative in a number of aspects, particularly in the way in which the movement of fish is modeled. Also, a number of simplifying assumptions were made about the basic population processes, which led to closed-form solutions for parameters of management interest, such as Maximum Sustainable Yield (MSY)- related quantities. The later, coupled with some newly-derived analytical results on decision theory, was used to derive optimal risk-averse fishing mortality targets and estimate them within the model framework.

In this report, I consider each of the model components separately, following the approach used during the review in Seattle. In each case, I first present a brief description of the modeling approach, followed by a critic. Recommendations are made when appropriate. Emphasis is placed on the modeling of the spatial dynamics. The other components, specifically, the assumptions about growth, recruitment and selectivity, may be problematic in some cases, but the potential problems are more obvious and so little time was devoted to them during the review. The concerns expressed here were discussed in Seattle. In some cases, alternatives for to how to evaluate (*e.g.* via simulations) the relevance and impact of the suspected problems were discussed. Dr. Thompson was able to develop and implement some of the simulations proposed, the results of which are summarized here.

Population dynamics

Growth. Mean weight at age is assumed to be a linear function of age. This simplification allows for closed-form expressions for total spawning biomass and exploitable biomass, which are used in the computation of MSY. As the model has an explicit age structure, however, alternative growth models could be used instead if the closed-form expressions were replaced by standard summations of age-specific terms. This would probably not entail too many extra computations. So, if the growth model turned out to be unrealistic for some stocks, it could be replaced. However, new equations to link the estimated parameters (F_{MSY}) to the stock-recruitment slope would need to be derived. The trade-offs between numerical efficiency and model performance could be easily evaluated within the ADMB framework.

Recommendation. Evaluate the suitability of the growth assumptions for each specific application.

Recruitment. Recruitment is represented as a linear (with an intercept) function of spawning biomass. This, together with the growth assumption, results in linear equations to represent that population dynamics processes and closed-form expressions used to parameterize the model in terms of F_{MSY} and B_{MSY}. As with the growth formulation, the suitability and effects of this assumption would need to be evaluated. The author has examined the implications of this assumption in the estimation of reference points using simpler models. The results were only discussed verbally and were not available for this review. It should be noticed that results derived using simple models should be applicable to the spatially structured model, as the stock-recruitment relationship does not depend on the spatial distribution of spawners and recruits.

In the present formulation, the variance of recruitment deviations is computed using the same coefficient of variation used to model process variability for other age groups. This is controlled by a single estimated parameter. The addition of one more parameter (or two if serial correlations in the recruitment time series were considered likely) would be recommended in order to capture the generally larger variability that characterizes the recruitment process compared to variability that affects older age classes. This can be very easily accomplished without changing the structure of the code.

Recommendation. Continue to explore the implications of the stock-recruitment assumptions. Add additional parameters to improve modeling of recruitment variability.

Selectivity. The model assumes that the selectivity is knife-edge, which is a standard assumption in simple age-aggregated models but not in age-structured models. This is a critical assumption in the model. Obviously, this assumption would conflict in cases when there are trends in selectivity over time, similarly to separable models. This, again, should be examined in each specific case.

Recommendation. Evaluate the suitability of the assumptions for each specific application.

Spatial dynamics

The model uses an empirical approach based on a random walk model to represent changes in the spatial distribution of fish over time. Changes in spatial distribution are estimated based on survey and commercial CPUE observations, as constrained by the stock dynamics equations. The advantage of the random-walk approach is that movement among cells does not need to be parameterized, which is usually a difficult task and likely very case-specific. A number of issues and concerns related to the spatial approach were raised during the review. Dr. Thomson was able to implement some initial simulation-estimation trials so that progress was made in evaluating some of the potential problems numerically.

Problems derived from nonrandom distribution of commercial CPUE observations

Information on changes on relative distribution of fish in space over time is obtained from spatially explicit survey and commercial CPUE observations. While, by design, the survey should cover the full spatial extent of the stock, the allocation of commercial CPUE observations depends on the fleet targeting practices and thus reflects the tendency of fishing effort to concentrate on areas of high fish abundance. On the one hand, this is good, as it means that directional displacements of the fleet can be very informative about fish movements (*e.g*. the case of seasonal migrations). On the other hand, in terms of a spatially-explicit approach, this implies that spatial cells with higher abundance tend to be preferentially sampled, while there may be fewer or no observations for lower-abundance cells. This would bias the estimates of spatial distribution and movement. When fish move, especially if they move directionally, the forward estimation equations will "see" fish concentrations appearing in different cells as fish move over time, but they will not "see" fish concentrations disappearing from previously observed cells. This is actually more complicated in the estimation approach developed by Dr. Thompson due to the effect of filtering. However, the basic problem remains. I anticipate that the effect would be to tend to smear the concentrations of fish in space and over time, especially when fish movement is directional and preferential targeting is strong.

Dr. Thompson conducted some simple simulations to examine this problem. He developed a simple algorithm to control the degree of directionality in the fish movement and the degree with which fishing effort was preferentially targeting areas of high abundance (as opposed to being randomly distributed). When both fish movement and effort allocation were random, the estimation model performed adequately in terms of estimating the changes in spatial distribution over time. The problem was most severe when the fish had a very strong directional component in their movement and effort was distributed in proportion to abundance. In that case, the estimation was able to identify the hot spots but the estimated high-abundance patches were maintained before and after they were actually formed. This is because there is information to point to areas of high abundance, but there is no data to show that abundance has come down.

So, while only a few simulations based on very small problems were examined, results are consistent with intuition. The severity of the problem will likely depend on the degree of directionality in the movement and the correlation of effort distribution with fish density.

It should be noted that this problem is not particular to the specific formulation of the spatial model reviewed but it results from well-known deficiencies of commercial fishery data. Other model formulations of the movement dynamics, however, may be more robust. For example, the more familiar movement models based on transition matrices may perform better, as they guarantee that fish are conserved through movement. In other words, for fish to appear in one cell they have to disappear from others. To some extent, this effect is present in Thompson's model through the fitting of total (aggregate) biomass data. This, however, does not impose a tight constraint in the model equations.

Recommendation: evaluate the extent of the problem using simulations tailored to the specific case.

Trade-offs associated with modeling the population dynamics in a spatially-aggregated versus disaggregated mode

In its present formulation, the population dynamic processes (recruitment, and fishing and natural mortality rates) are not a function of spatial location. In particular, fishing mortality is a single aggregate parameter (for each year), which is applied to all spatial cells. Most data used in the fit are aggregate (catch in biomass, commercial and survey catch at age and total biomass from the survey), except for commercial and survey CPUE, which are spatially explicit. In that case, what are the advantages of *simultaneously* estimating the stock dynamics (which in the present formulation is not a function of space) and the spatial distribution (which is not a function of age)? This question was raised during the review in terms of two aspects:

1. Is the aggregate stock assessment improved by the incorporation of the spatial dynamics? Is it degraded?

2. Is the assessment of the spatial distribution of fish over time improved by inclusion of the age-structure aggregate stock dynamics?

Regarding the first question, the inclusion of the spatial dynamics allows for the use of commercial and survey CPUE data. The survey catch rate, however, is still used in the absence of spatial structure in the form of a survey biomass index. So it is in fact just the commercial CPUE data that would be lost by collapsing the spatial dimension. Given the considerations raised above, the value of these data (in terms of estimating aggregate abundance) will depend on the degree of preferential allocation of effort on areas of high abundance. If effort is random, then the use of CPUE is beneficial. However, in that case, average CPUE (or effort) could be used as an aggregate index and so, again, the aggregate assessment would not necessarily benefit from carrying the spatial structure simultaneously. In the likely more common situation in which CPUE is not randomly distributed, its inclusion (under the assumption of constant catchability) may actually degrade the aggregate assessment.

In conclusion, while the issue would need to be explored with simulations, it is my impression that, *as long as the dynamics are not location-dependent*, the aggregate stock assessment will not benefit from modeling the spatial structure. Of course if catch and effort data were made available at a smaller spatial resolution, which is very likely given the priority placed on spatially-related issues, then fishing mortality parameters could be made a function of location and then the benefits of doing a spatially explicit assessment would be clear. The model would need to be extended to accommodate this new type of data.

The second question is important given the need to address a number of issues related to the spatial distribution of fish and effort, beyond the aggregate stock assessment. This was in fact what motivated the development of this model. The alternative to modeling the age structure and the spatial structure simultaneously is to consider the changes in the relative distribution of fish over time and space independently of the stock dynamics using an empirical model (such as the one used in this model). However, as discussed earlier, the stock dynamics is needed to constrain the total abundance of fish to assure that fish are conserved through the movement. This could still be accomplished by imposing a constraint or prior distribution on abundance trends derived from the assessment model independently.

In conclusion, the benefits of conducting a spatially-explicit stock assessment would be clearer if the population dynamics equations were a function of space, and catch and effort data were available at the model spatial resolution to enable estimation of location-dependent model parameters.

Recommendation: consider uncoupling the two components (age-specific dynamics and spatial dynamics) until information on catches and effort is available at a finer spatial resolution to justify the inclusion of dynamic parameters that vary with location. Evaluate pros and cons using simulations.

Parameter estimation

The model estimates a few independent parameters (fishing mortalities per year, variances for the dynamic processes and the observations -one parameter for the dynamics, two for the random movement in space- biomass and fishing mortality at MSY, commercial and survey catchabilities and natural mortality). Other parameters are estimated explicitly conditioned on the rest, and the stochastic variables are estimated via a Kalman filter. The estimation code has been implemented using AD Model Builder (ADMB), which allows efficient parameter estimation for nonlinear models.

Correctness of derivations and computing code

Dr. Thompson has carefully checked the equations and algebra showing that closed-form solutions and alternative equations produce similar numerical results. The estimation code has been debugged and cross-checked by comparing numerical results obtained with ADMB and the likelihood values computed using MATHCAD.

Efficiency considerations

Numerical efficiency will need to be improved to facilitate testing of the estimation model by simulations. Simulations done so far, which have mostly involved reduced dimensions, suggest that some effort spent in that direction would be guaranteed.

The C++ compiler and ADMB have tools that can be used to identify the most computing-intensive parts of the objective function and the amount that is being written to disk. Those tools could be used to help decide where to spend more effort in order to save memory or speed up computations.

In the absence of those diagnostics, I examined the formulation and code to try to identify places where savings could be made. Results were as follows:

- A bottleneck identified by Dr. Thomson is a large matrix (number of spatial cells × number of ages) invertion done at each time step to specify the process variance as a function of the stationary distribution of the state variables. An approach to reduce computations was identified during the review: because the movement model is purely random, the stationary distribution of abundance in space is simply proportional to the area of the cells (as the author was able to confirm numerically). In that case it is only the abundance at age for the aggregated population that needs to be solved for, which can be done much more efficiently. This modification was implemented during the review and resulted in some appreciable gains in efficiency.
- Kalman equations are specified defining an observation equation that involves a "slope" matrix Ω (declared as a variable object in ADMB). This matrix can be decomposed into a constant matrix (except for the weight of the plus group) and a diagonal matrix that contains all the variables that depend on estimated parameters (fishing mortalities and catchabilities). The constant matrix does the partial summations and cross-products needed to get total abundances at age (summed over space), total biomass (summed over space and ages) and total abundance per cell

(summed over ages). These partial summations could be computed more efficiently using simple summations and cross products (instead of full-matrix multiplications) to get the expected value of the needed aggregate state variables and their covariance matrices before they are multiplied (element-wise) by the variable parameters (the diagonal matrix). Saving would be even more substantial in the derivative computations involved, which are invisible to the user.

- Similarly, the use of matrix multiplication for the process equations, while it provides a compact notation for paper presentation, is not the most efficient way to compute expected state transitions and covariances, as most of the entries of the transition matrix (Θ) are zeros. Thus, standard recursive equations for abundance at age-cell (i.e. $N_{t+1,a+1,c} = f(N_{t,a,c})$) should be used instead.
- The implementation of the model in ADMB has minimized the number of independent parameters to be estimated at the expense of increasing the amount of computations involved per function evaluation. This was achieved through a smart use of model properties by solving for some of the parameters explicitly as conditional functions of the independent parameters. Specifically, this was done with the parameters related to the initial conditions (*i.e.* abundance at age and spatial cell at the start of the time series). Trade-offs between increasing the number of independent parameters and simplifying the objective function could be evaluated.

Recommendation: replace matrix equations for the observation and population dynamic transitions by standard case-by-case observation and transition equations. Use available ADMB and C++ compiler tools to identify places where alternative, more efficient formulations would be most beneficial. Optimize the use of memory to minimize writing to disk.

GENERAL ISSUES ABOUT SPATIAL RESOLUTION AND DATA AVAILABILITY

Some of the problems discussed above are not restricted to Dr. Thompson's model. Rather, they point to some intrinsic limitations in our capacity to answer questions pertaining to the spatial distribution of fish stocks and the local effects of fishing beyond the resolution of the data currently available. Emphasis should be placed on the collection of data at the appropriate spatial resolution, which will depend on the ecological or management question to be addressed.

RISK-AVERSE FISHING TARGETS FOR FISHERIES MANAGEMENT

A major focus in fishery research has been the identification of management approaches that are robust to uncertainties in the assessments and future trends in stock abundance. A number of generic decision rules have been derived to guide management decisions, which are intended to perform well under a variety of scenarios. The alternative to the use of generic decision rules is to design management strategies for each specific case, which can better account for the specific nature of the fishery, the uncertainty in the stock assessment and the quality of the information available to monitor future stock status. In my view, the latter is the preferred approach. However, generic rules can be useful to guide management decisions when specific policy evaluations are not available.

Dr. Thompson has produced a very important contribution in this area using decision theory to derive risk-averse fishing targets that account for process variability and uncertainty in the estimates of current stock status. His results have been applied to the management of groundfish in the North Pacific([1]). He has now expanded these results to incorporate also the effects of parameter estimation uncertainty in the computation if risk-averse fishing targets. The results provide quantitative rules for decreasing fishing targets as a function of the variance of the stationary distribution of relative yield, itself a function of the three sources of uncertainty considered. Equations have been derived to solve for this variance parameter using Kalman filter estimates and the ability of AD Model Builder to approximate variances of derived parameters. While the decision analysis frame is general, the application is based on the specific model equations. Results are important and potentially very useful; it would be interesting to investigate their applicability beyond the context of the specific model reviewed.

In order to convey better the implications of following the proposed optimal rules (both for the stock and the fishing industry), Monte Carlo simulations could be conducted to evaluate them in terms of standard performance statistics such as future yields, effort, future stock trends, etc. A difficulty with the decision analysis is that optimality is defined in terms of maximizing certain risk-averse utility functions, which meaning tends to be obscure.

Recommendation: Evaluate the performance of the derived optimal decision rules in terms of parameters of management interest.

---

[1] Environmental assessment/regulatory impact review for Amendment 44 to the Fishery Management Plan for the groundfish fishery of the Bering Sea and Aleutian Islands area and Amendment 44 to the fishery management plan for the groundfish fishery of the Gulf of Alaska to redefine Acceptable Biological Catch and Overfishing. June 18, 1996.

APPENDICES

1. STATEMENT OF WORK

Consulting Agreement Between The University of Miami and Dr. Ana Parma

September 29, 2011

September 29, 2011

General

The Alaska Fisheries Science Center (AFSC) requests review of a modeling framework which is being proposed for future use in the Bering Sea / Aleutian Islands and Gulf of Alaska Pacific Cod assessments. This review would come at a pivotal point, as the Pacific cod assessments are two of the most complex assessments conducted by the Center and the proposed modeling framework would address several questions that have arisen regarding the current assessment models, the productivity and spatio-temporal distribution of the stocks, and the management of the fisheries. In addition, Pacific cod is an important component of the ecosystem both as predator and prey. For example, Pacific cod has been found to be an important prey item in the diet of Steller sea lions, an endangered species. These factors create a compelling need for independent peer review of the proposed modeling framework.

The consultant will need to be thoroughly familiar with various subject areas involved in the proposed modeling framework, including population dynamics, spatial processes, the Kalman filter, the AD Model Builder programming language, and decision theory. The consultant will travel to Seattle, Washington to discuss the proposed modeling framework with the lead analyst for the Pacific cod assessments. The report generated by the consultant should include:

a. A statement of the strengths and weaknesses of the proposed modeling framework;
b. Recommendations for alternative model configurations or formulations.

AFSC will provide copies of the most recent stock assessment documents; the most recent external review of the Pacific cod assessments; the current research plan of the Pacific Cod Working Group; a description of the proposed modeling framework, including a list of symbols used, derivations of the functional forms used, and a set of worked examples; and AD Model Builder code for the parameter estimation routine.

Specific

The consultant's duties shall not exceed a maximum of three weeks - several days for document review, a 5-day meeting, and several days to produce a written report of the findings.

The itemized tasks of the consultant include:

1. Read and become familiar with the relevant documents provided to the consultant;
2. Discuss the proposed modeling framework with the lead analyst in Seattle, Washington, from July 30 to August 3, 2001;
3. No later than August 27, 2001, submit a written report of findings, analysis, and conclusions. The report should be addressed to the "UM Independent System for Peer Reviews, " and sent to David Die, UM/RSMAS, 4600 Rickenbacker Causeway, Miami, FL 33149 (or via email to ddie@rsmas.miami.edu).

Signed________________________

Date____________

PRELIMINARY BUDGET

1. Salary ($600 per day for 21 days)	$12,600
2. Plane fare (Buenos Aires to Seattle)	$1,800
3. Lodging (July 29-August 4: 6 nights)	$900
4. Meals ($30 per diem for 7 days)	$210
5. Car rental	$300
6. Additional transportation	$200
TOTAL	$16,010

ANNEX I: REPORT GENERATION AND PROCEDURAL ITEMS

1. The report should be prefaced with an executive summary of findings and/or recommendations.

2. The main body of the report should consist of a background, description of review activities, summary of findings, and conclusions/recommendations.

3. The report should also include as separate appendices the bibliography of materials provided by the Center for Independent Experts and the center and a copy of the statement of work.

4. Individuals shall be provided with an electronic version of a bibliography of background materials sent to all reviewers. Other material provided directly by the center must be added to the bibliography that can be returned as an appendix to the final report.

Please refer to the following website for additional information on report generation: http://www.rsmas.miami.edu/groups/cimas/Report_Standard_Format.html